CLINICAL SCIENCE

# VF-14 item specific responses in patients undergoing first eye cataract surgery: can the length of the VF-14 be reduced?

**D S Friedman, J M Tielsch, S Vitale, E B Bass, O D Schein, E P Steinberg**

*Br J Ophthalmol* 2002;**86**:885–891

See end of article for authors' affiliations

Correspondence to: David S Friedman, MD, MPH, Dana Center for Preventive Ophthalmology, Wilmer 120, Johns Hopkins University, 600 North Wolfe Street, Baltimore, MD 21287, USA; David.Friedman@jhu.edu

Accepted for publication 15 February 2002

**Aims:** To report the item specific responses of the VF-14 in a population of patients undergoing cataract surgery in their first eye and to determine whether or not the VF-14 can be reduced without compromising its performance as an index of cataract related visual impairment.

**Methods:** The item specific responses to the VF-14 were analysed before (771 patients) and 4 months after (552 patients) cataract surgery in one eye to determine if the VF-14 index can be reduced without compromising its performance. Patients studied were selected from a cross sectional longitudinal study of patients undergoing cataract surgery in 72 ophthalmologist's offices located in three metropolitan regions of the United States.

**Results:** Pairwise correlations between items in the VF-14 were all less than 0.6, indicating that no items could be removed solely on the basis of redundancy. 10 items correlated moderately with change in trouble, and 11 correlated moderately with change in satisfaction ($r$ >0.15) at 4 months after cataract extraction. Eleven items demonstrated an effect size >0.4 at 4 months. These 11 items were either important for detecting cataract related functional disability or for quantifying the extent to which cataract impaired function. Additionally, 11 items were needed to detect adequately individuals with functional impairment. Three items (recognising people, cooking, and reading large print) were less responsive to cataract extraction and were more strongly associated with ocular comorbidities.

**Conclusions:** While previous reports indicate that the VF-14 can be significantly shortened, our analysis only justifies removing three items. While the resulting VF-11 has properties similar to the VF-14, the limited time savings do not justify altering this already validated instrument.

Cataract extraction remains the most commonly performed operation on Medicare beneficiaries, with 1.4 million surgeries in 1998 (Kevin Hayes, personal communication, Medicare Payment Advisory Commission, Washington, DC, 19 August 1999). To quantify the functional limitations associated with cataract, researchers have developed standardised questionnaires designed to measure the impact of impaired vision on patients' ability to perform daily activities.[1–4] Patients' responses have correlated well with patient reported trouble and satisfaction with vision before surgery and have improved significantly after cataract surgery.[1 5–7] In contrast, preoperative Snellen acuity testing does not assess the functional difficulties experienced by patients with cataract and has been documented to be poorly correlated with patient reported trouble and satisfaction with vision.[5–8]

Reports issued by the Agency for Health Care Policy and Research Cataract Guidelines Panel and the American Academy of Ophthalmology state that the goal of cataract surgery is functional improvement,[9 10] and the indication for cataract surgery is cataract induced functional impairment that is considered to be significant by the patient.

Despite the widespread availability of reliable, valid, and responsive instruments for quantifying functional impairment related to vision, most practising ophthalmologists still do not use them. Possible reasons for this low utilisation include lack of belief in the utility of the available instruments, low awareness of them among ophthalmologists, and a reluctance to spend the time required to administer a standardised questionnaire to a patient. The purpose of this study was to determine whether the VF-14 could be reduced in length without substantial compromise of its performance. Expanded use of standardised methods of obtaining this information would improve the assessment of cataract patients in clinical practice and evaluation of the care they receive in different settings.[1]

A recent publication indicated that the VF-14 might be able to be reduced to seven questions without compromising performance.[11] In order to further examine these findings, we sought to determine which of the individual items within the VF-14 were most responsive to cataract surgery among the population of patients enrolled in the Cataract Patient Outcomes Research Team (PORT) study and to assess the correlation between alternative combinations of these items and self reported trouble and satisfaction with vision.[5 12]

## METHODS

Data collected during the Cataract PORT longitudinal study of cataract outcomes were used for this analysis. The methods employed in that observational study of cataract surgery have been described in detail elsewhere.[1 8] In brief, patients of age 50 years or older scheduled to undergo first eye cataract surgery by any of 72 ophthalmologists in three metropolitan areas were recruited between July and December 1991 for participation. Data were collected preoperatively and at 48 hours, 4 months, and 12 months after cataract extraction. The VF-14, which asked patients about 14 vision dependent activities, was used to assess functional impairment related to vision.[1] The VF-14 is administered by asking patients the following: "Do you have any difficulty, even with glasses, in doing any of the following activities?" If the answer to a question regarding a particular activity is "yes," the patient is asked whether his/her level of difficulty with performance of the activity is "a little," "a moderate amount," "a great deal," or whether he/she is unable to perform the activity because of

his/her vision. In addition, the patient can state that the activity is "not applicable" (that is, he/she does not perform the activity for reasons unrelated to his/her vision). The activities addressed by the VF-14 are:

Reading small print such as labels on medicine bottles, or a telephone book

Reading a newspaper or book

Reading a large print book or large print newspaper or numbers on a telephone

Recognising people when they are close to you

Seeing steps, stairs, or curbs

Reading traffic, street, or store signs

Doing fine handwork like sewing, knitting, crocheting, or carpentry

Writing cheques or filling out forms

Playing games such as bingo, dominoes, card games, or mahjong

Taking part in sports like bowling, handball, tennis, or golf

Cooking

Watching television

Daytime driving

Night-time driving

The VF-14 is scaled from 0 to 100, with 0 indicating that the patient is unable to perform any applicable activities and 100 meaning that the patient can perform all applicable activities without difficulty.

Patients enrolled in the Cataract PORT study also were asked to rate their trouble with vision and their satisfaction with vision, using scales with four possible responses. For trouble with vision, response options were "none," "a little," "a moderate amount," or "a great deal." For satisfaction with vision, response options were "very dissatisfied," "dissatisfied," "satisfied," or "very satisfied."

A total of 888 patients were referred for enrolment during the Cataract PORT study; 772 (86.9%) agreed to participate. Baseline analyses were performed on 771 individuals who completed the VF-14. Analyses of the effect of cataract surgery on patient reported outcomes were limited to the 552 patients who had both preoperative and 4 month postoperative VF-14 data and had cataract surgery in only one eye.

### Statistical analysis

Differences in the proportion of patients enrolled in the Cataract PORT study who reported particular levels of difficulty with specific VF-14 items were assessed using a $\chi^2$ test. Associations between preoperative visual acuity, ocular comorbidities, and item specific responses were assessed using multiple logistic regression. Preoperative versus postoperative changes in proportions for paired data were assessed using McNemar's test. Associations between preoperative versus postoperative change in reported trouble with vision and satisfaction with vision and preoperative versus postoperative responses to each item in the VF-14 index were evaluated using Spearman's rank correlation coefficient. Effect sizes were calculated for each item in the VF-14 as well as for the VF-14 and indices we constructed that included fewer than the original 14 items. The effect size is a measure that can be used to quantify changes in health status for a group produced by an intervention. While there are several approaches to measuring effect size,[13 14] we used the mean change in the item or index score following cataract surgery divided by the standard deviation of that item or index score at baseline. An effect size of one, therefore, means that the score, on average, changed by one standard deviation. The effect size for the VF-14 (0.99) has been reported previously.[12] An instrument demonstrating an effect size greater than 0.8 is considered to be responsive.[15]

We explored the impact of removal of an item from the VF-14 if it did not appear to detect uniquely a common type of functional impairment. Specifically, we assessed the impact of removal of an item if few Cataract PORT study enrollees reported that the activity applied to them, if over 90% of subjects reported no difficulty with the activity at baseline, if the item correlated poorly with baseline trouble and satisfaction with vision, or if the item did not uniquely identify a functional impairment for a given patient. We also assessed the association between patient responses to each item and an index of ocular comorbidity (comorbidity was present if the subject had macular degeneration, glaucoma, or diabetic retinopathy in the opinion of the examining ophthalmologist) to determine if conditions other than cataract may have been responsible for patient reported functional impairment. Items that were strongly associated with ocular comorbidities were considered as candidates for removal from the VF-14. A final component of our strategy for paring down the VF-14 was to remove items that did not appear to be responsive to cataract surgery, as demonstrated by a small effect size, and/or poor correlation with patient reported change in trouble and satisfaction with vision after cataract removal. Factor analysis on the final, reduced index was performed to determine whether there were subscales within the overall scale that reflected clinically meaningful domains of function. Cronbach's $\alpha$[16] was also calculated to measure the internal consistency of shortened forms of the VF-14.

## RESULTS

### Patient characteristics

Subjects enrolled in the Cataract PORT study were mostly white (94%) and female (63%, Table 1). The 552 subjects who underwent surgery in only one eye by 4 months were similar to those who had both eyes in age, race, and sex, but had

**Table 1** Baseline characteristics of all enrolled patients, those who had only one eye surgery at 4 months, and those who underwent surgery in both eyes by 4 months

| | N = 766* | N = 552 | N = 219 |
|---|---|---|---|
| Age (years, SD) | 72 (8) | 72 (8) | 73 (7) |
| Sex (female) | 63% | 63% | 63% |
| Race (white) | 94% | 94% | 94% |
| VF-14 at baseline (mean, SD) | 75.5 (17) | 76.1 (17) | 72.5 (18)† |
| Median visual acuity | | | |
| Better eye | 20/40 (20/20–HM) | 20/40 (20/20–HM) | 20/50 (20/20–HM) |
| Worse eye | 20/70 (20/20–NLP) | 20/60 (20/20–NLP) | 20/70 (20/30–LP) |

*Both preoperative and clinical data available for 766 of 771 subjects.
†Statistically significantly different from those with surgery for one eye adjusting for age, race, and sex, p<0.01
LP = light perception, NLP = no light perception, HM = hand movements.
Better eye operated on in 7% of patients.

**Table 2** Preoperative responses to individual VF-14 items at baseline (n = 771)

| Activity | Applicable | No difficulty | Little | Moderate | Great deal or unable |
|---|---|---|---|---|---|
| Small print | 769 (99.7%) | 146 (19.0%) | 77 (10.0%) | 118 (15.3%) | 427 (55.5%) |
| Newspaper | 765 (99.2%) | 303 (39.6%) | 68 (8.9%) | 148 (19.3%) | 246 (32.2%) |
| Large print | 747 (96.9%) | 682 (91.3%) | 15 (2.0%) | 28 (3.7%) | 22 (2.9%) |
| Recognise people | 771 (100%) | 698 (90.5%) | 21 (2.7%) | 21 (2.7%) | 31 (4.0%) |
| Steps | 767 (99.5%) | 544 (70.9%) | 59 (7.7%) | 91 (11.9%) | 73 (9.5%) |
| Signs | 761 (98.7%) | 450 (59.1%) | 51 (6.7%) | 114 (15.0%) | 146 (19.2%) |
| Sewing | 583 (75.6%) | 237 (40.7%) | 46 (7.7%) | 92 (15.8%) | 209 (35.8%) |
| Checks | 741 (96.1%) | 555 (74.9%) | 41 (5.5%) | 68 (9.2%) | 77 (10.4%) |
| Bingo | 500 (64.9%) | 404 (80.8%) | 18 (3.6%) | 47 (9.4%) | 31 (6.2%) |
| Sports | 195 (25.3%) | 139 (71.3%) | 12 (6.2%) | 22 (11.3%) | 22 (11.3%) |
| Cooking | 654 (84.8%) | 587 (89.8%) | 11 (1.7%) | 33 (5.0%) | 23 (3.5%) |
| Television | 761 (98.7%) | 522 (68.6%) | 79 (10.4%) | 92 (12.1%) | 68 (8.9%) |
| Day driving | 619 (80.3%) | 310 (50.1%) | 141 (22.6%) | 97 (15.7%) | 72 (11.6%) |
| Night driving | 619 (80.3%) | 53 (8.6%) | 89 (14.4%) | 96 (15.5%) | 381 (61.6%) |

higher baseline scores on the VF-14 ($p < 0.01$, adjusted for age, race, and sex).

### Performance of individual items of the VF-14

The proportion of study enrollees for whom individual VF-14 activities were applicable among the 771 individuals with baseline data ranged from 25.3% for taking part in sports to 100% for recognising people (Table 2). The proportion of respondents for whom specific VF-14 activities were applicable at both the preoperative and postoperative evaluations ranged from 16.7% for taking part in sports to 99.8% for reading small print and recognising people (Table 3). Over 300 patients (55%) participated in each activity except for sports at both baseline and 4 months postoperatively.

For both the entire enrolled cohort (n = 771) and the 552 patients with baseline and 4 month postoperative data and no second eye surgery before the 4 month postoperative examination, 10% of patients or fewer reported any difficulty at baseline in recognising people, reading large print, or cooking. In contrast, over half reported at baseline that they either had a great deal of difficulty or were unable to drive at night or read small print. The proportion of patients who reported moderate or severe impairment with an activity decreased significantly for each activity after cataract extraction ($p < 0.001$ for all activities).

Pairwise correlations between the scores on individual items in the VF-14 were all <0.6, indicating that variables could not be removed from the VF-14 on the basis of redundancy alone. Correlations between the score on each item and the score on the remaining 13 items were lowest for difficulty reading large print, recognising people, cooking, and night driving (<0.4) suggesting that these items were measuring impairments that were the most different from the other 10 items. These differences could be due to these activities identifying types of cataract related functional limitations not captured by the other items, or to these activities being more affected by non-cataract ocular disorders. Night driving may be particularly impaired by cataract induced glare. Hence, it is possible that patients may have considerable difficulty with night driving without having much difficulty with other activities included in the VF-14. In contrast, the comparatively low correlation between scores for reading large print, recognising people, or cooking and the other items may be due to the fact that the vast majority of individuals had "no difficulty" performing these three activities at baseline.

To assess quantitatively the unique contribution of individual items to the VF-14 score, as well as the responsiveness of individual questions to cataract surgery, we examined, for each item: (1) the correlation between the change in the item score and the change in trouble and satisfaction with vision after cataract extraction; (2) the effect size of the item; (3) the effect size of the remaining 13 item index score after removing that item from the VF-14; (4) the frequency with which a patient reported any difficulty in that activity; and (5) the

**Table 3** Preoperative and 4 months postoperative responses to individual VF-14 items among patients for whom a given activity was applicable at baseline and 4 months (n = 552)*

| | | Reported Level of Difficulty with Activity | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Preoperative (%) | | | | 4 Months postoperative (%) | | | |
| Activity | Number for whom activity was applicable (%)† | None | Little | Mod. | Great deal or unable | None | Little | Mod. | Great deal or unable |
| Reading small print | 551 (99.8) | 21.1 | 10.2 | 15.6 | 53.2 | 61.0 | 15.8 | 11.3 | 12.0 |
| Reading the newspaper | 543 (98.4) | 41.8 | 8.7 | 19.9 | 29.7 | 87.7 | 4.4 | 3.1 | 4.8 |
| Reading large print | 528 (95.7) | 90.5 | 2.1 | 4.4 | 3.0 | 98.7 | 0.2 | 0.6 | 0.6 |
| Recognising people | 551 (99.8) | 91.1 | 2.5 | 2.7 | 3.6 | 98.4 | 0.7 | 0.5 | 0.4 |
| Seeing steps or curbs | 548 (99.3) | 71.4 | 6.4 | 12.8 | 9.5 | 89.6 | 6.2 | 2.9 | 1.3 |
| Seeing signs | 541 (98.0) | 62.1 | 6.1 | 15.3 | 16.5 | 91.3 | 3.5 | 2.6 | 2.6 |
| Doing fine handwork | 345 (62.5) | 39.7 | 9.0 | 16.0 | 35.4 | 88.1 | 4.1 | 3.5 | 4.4 |
| Writing checks | 523 (94.7) | 75.9 | 5.0 | 10.1 | 9.0 | 95.8 | 1.9 | 1.0 | 1.3 |
| Playing games | 305 (55.3) | 79.7 | 3.3 | 12.1 | 4.9 | 97.7 | 1.3 | 0.3 | 0.7 |
| Playing sports | 92 (16.7) | 72.8 | 3.3 | 14.1 | 9.8 | 95.7 | 1.1 | 1.1 | 2.2 |
| Cooking | 452 (81.9) | 90.0 | 2.0 | 4.9 | 3.1 | 98.7 | 0.2 | 0.7 | 0.4 |
| Watching TV | 546 (98.9) | 70.2 | 10.3 | 12.8 | 6.8 | 95.4 | 2.2 | 1.8 | 0.6 |
| Daytime driving | 411 (74.5) | 52.3 | 23.6 | 16.8 | 7.3 | 90.3 | 6.8 | 2.0 | 1.0 |
| Night driving | 411 (74.5) | 8.5 | 17.0 | 17.0 | 57.4 | 42.1 | 17.3 | 6.3 | 34.3 |

*Average score for each item improved, $p < 0.001$
†% of all patients for whom activity was applicable at both baseline and 4 months post-op

**Table 4** Correlation between postoperative change in the response for each activity and change in trouble and satisfaction with vision

| Activity | Correlation with change in trouble* | Correlation with change in satisfaction* |
|---|---|---|
| Reading small print | 0.26 (0.04) | 0.27 (0.04) |
| Reading the newspaper | 0.29 (0.04) | 0.24 (0.04) |
| Reading large print | 0.11 (0.05) | 0.05 (0.05) |
| Recognising people | 0.09 (0.04) | 0.11 (0.04) |
| Seeing steps or curbs | 0.14 (0.04) | 0.15 (0.04) |
| Seeing signs | 0.16 (0.04) | 0.16 (0.04) |
| Doing fine handwork | 0.33 (0.05) | 0.26 (0.05) |
| Writing checks | 0.20 (0.04) | 0.18 (0.04) |
| Playing games | 0.17 (0.06) | 0.16 (0.06) |
| Playing sports | 0.29 (0.10) | 0.21 (0.10) |
| Cooking | 0.09 (0.05) | 0.18 (0.04) |
| Watching TV | 0.16 (0.04) | 0.07 (0.04) |
| Daytime driving | 0.24 (0.05) | 0.16 (0.05) |
| Night driving | 0.19 (0.05) | 0.17 (0.05) |

*Spearman correlation coefficients (SD).

**Table 6** Impact of removal of an activity from VF-14 on the effect size*

| Activity | % with no difficulty at baseline | Effect size of index (VF-13) if activity were removed from the VF-14 |
|---|---|---|
| Reading small print | 21.1 | 0.98 |
| Reading the newspaper | 41.8 | 1.02 |
| Reading large print | 90.5 | 1.07 |
| Recognising people | 91.1 | 1.06 |
| Seeing steps or curbs | 71.4 | 1.08 |
| Seeing signs | 62.1 | 1.04 |
| Doing fine handwork | 75.9 | 1.08 |
| Writing checks | 39.8 | 0.97 |
| Playing games | 79.7 | 1.06 |
| Playing sports | 72.5 | 1.07 |
| Cooking | 90 | 1.06 |
| Watching TV | 70.2 | 1.04 |
| Daytime driving | 54.4 | 1.04 |
| Night driving | 9 | 0.99 |

*Effect size for the VF-14 is 0.99.

likelihood that a patient would have difficulty with at least one item in the remaining VF-13 after an item was removed from the VF-14.

Postoperative changes in ability to read large print, recognise people, and see steps or curbs were the most weakly correlated with both change in trouble and change in satisfaction with vision ($r \leqslant 0.15$ for both trouble and satisfaction, Table 4). The items in which the change in score was most strongly correlated with change in trouble and satisfaction ($r \geqslant 0.25$) were reading small print and doing fine handwork. The effect size of individual questions (how many baseline standard deviations a measure changed after cataract extraction) ranged from 1.0 for reading small print to 0.21 for recognising people (Table 5). The activities that were most responsive to surgery (as measured by effect size $\geqslant 0.8$) were near vision activities (reading small print, reading the newspaper, and doing fine handwork). The activities that were least responsive to surgery (as evidenced by an effect size $\leqslant 0.3$) were recognising people, reading large print, and cooking. The poor responsiveness of these items is likely due to the large number of patients who could not improve on these activities given the lack of difficulty they had with them at baseline. The effect sizes of the various VF-13 indices that result from removing items one at a time were similar to those found for the full VF-14.

Some questions included in the VF-14 provided little information about the functional status of most patients. For example, four out of five patients did not participate in sports either preoperatively or postoperatively, meaning that, for the majority of patients, this question did not contribute to their VF-14 score. In addition, some items (reading large print, recognising people, and cooking) asked about activities with which $\geqslant 90\%$ of individuals had no difficulty. To determine quantitatively which items contributed unique information we assessed (using data from all subjects at baseline) the frequency that an individual activity was the only one for which a patient reported having difficulty. One could anticipate, for example, that a patient might have trouble playing sports as her or his sole complaint. In this analysis, for each activity we used the number of patients reporting moderate or greater difficulty preoperatively as the denominator. For these patients, the mean number of other questions for which they reported moderate or greater difficulty was calculated. This mean number ranged from 2.3 for night driving to 5.9 for playing games (Table 6). In addition, for each activity, of the individuals reporting moderate or greater difficulty, the percentage reporting moderate or greater difficulty on less than two (that is, none or one) other activities was calculated. Of the 771 individuals in our analysis, 26.1% reported moderate or greater difficulty on driving at night but had either zero or only one other activity for which they reported moderate or greater difficulty. Over 15% of those enrolled reported difficulty reading small print and had either one or no other items with which they had moderate or worse difficulty. Therefore, the reading small print and driving at night items

**Table 5** Effect size* of individual questions in the VF-14

| Function | Preoperative Mean (SD) | 4 months Mean (SD) | Change | Effect size |
|---|---|---|---|---|
| Reading small print | 1.89 (1.33) | 3.22 (1.15) | 1.33 | 1.00 |
| Reading the newspaper | 2.59 (1.34) | 3.74 (0.77) | 1.15 | 0.86 |
| Reading large print | 3.78 (0.77) | 3.96 (0.34) | 0.18 | 0.23 |
| Recognising people | 3.81 (0.66) | 3.95 (0.46) | 0.14 | 0.21 |
| Seeing steps or curbs | 3.39 (1.38) | 3.84 (0.52) | 0.45 | 0.33 |
| Seeing signs | 3.12 (1.21) | 3.84 (0.60) | 0.72 | 0.60 |
| Doing fine handwork | 2.40 (1.47) | 3.74 (0.85) | 1.34 | 0.91 |
| Writing checks | 3.50 (1.04) | 3.92 (0.42) | 0.47 | 0.45 |
| Playing games | 3.59 (0.91) | 3.97 (0.26) | 0.38 | 0.42 |
| Playing sports | 3.38 (1.15) | 3.86 (0.84) | 0.48 | 0.42 |
| Cooking | 3.79 (0.68) | 3.97 (0.25) | 0.18 | 0.26 |
| Watching TV | 3.44 (0.95) | 3.93 (0.37) | 0.49 | 0.52 |
| Daytime driving | 3.11 (1.16) | 3.84 (0.63) | 0.73 | 0.63 |
| Night driving | 1.33 (1.40) | 2.35 (1.78) | 1.02 | 0.73 |

*Effect size is the mean change in the value divided by the standard deviation of values at baseline.

**Table 7** Resulting number of individuals identified as having difficulty after removing one item from the VF-14

| | Number identified as having a little difficulty or worse on at least one item | Number identified as having moderate difficulty or worse on at least one item | Number identified as having a great deal of difficulty or worse on at least one item |
|---|---|---|---|
| Total identified with VF-14 | 722 | 642 | 438 |
| Activity removed from the VF-14 | | | |
| Reading small print | 705 | 605 | 417 |
| Reading the newspaper | 716 | 639 | 436 |
| Reading large print | 722 | 642 | 438 |
| Recognising people | 722 | 642 | 437 |
| Seeing steps or curbs | 722 | 642 | 438 |
| Seeing signs | 720 | 639 | 437 |
| Doing fine handwork | 719 | 635 | 428 |
| Writing checks | 722 | 641 | 438 |
| Playing games | 722 | 642 | 438 |
| Playing sports | 720 | 645 | 438 |
| Cooking | 722 | 642 | 438 |
| Watching TV | 721 | 642 | 438 |
| Daytime driving | 720 | 639 | 436 |
| Night driving | 720 | 530 | 182 |

often identified individuals with visual limitations who might have been missed if these questions were removed. Conversely, every patient who reported moderate or greater difficulty reading large print or watching TV had at least two other activities for which he/she reported moderate or greater difficulty. Less than 1% of the study population reported moderate or worse difficulty preoperatively reading large print, recognising people, seeing steps or curbs, playing games, playing sports, cooking, or watching television who did not have at least two other items on which they reported moderate or worse difficulty.

Of the 771 individuals enrolled in the Cataract PORT study, 722 had at least one activity in the VF-14 with which they had at least a little difficulty (all patients had already elected to have cataract surgery). Removing reading small print led to the greatest decline in the number of individuals identified as having at least one question with moderate or worse difficulty (Table 7), indicating that for many patients this item captures functional difficulty that is not identified by other items in the VF-14 index. Night driving was the only other item with a similar profile.

**Table 8** Association of ocular comorbidity* with having moderate or worse difficulty with items on the VF-14

| Item | Odds ratio† (95% CI) |
|---|---|
| Small print | 1.2 (0.7 to 2.1) |
| Newspaper | 1.2 (0.9 to 1.6) |
| Large print | 1.4 (0.8 to 2.7) |
| Recognise people | 2.9 (1.6 to 5.2) |
| Steps | 1.6 (0.95 to 2.0) |
| Signs | 1.4 (0.9 to 2.1) |
| Fine handwork | 1.0 (0.7 to 1.5) |
| Cheques | 1.4 (0.9 to 2.2) |
| Bingo | 1.3 (0.7 to 2.2) |
| Sports | 1.3 (0.6 to 2.7) |
| Cooking | 1.9 (1.1 to 3.5) |
| Television | 1.2 (0.8 to 1.9) |
| Day driving | 1.5 (0.97 to 2.3) |
| Night driving | 1.2 (0.7 to 2.0) |

*Comorbidity defined by the clinician as the presence of glaucoma, macular degeneration, or diabetic retinopathy.
†Odds ratio represents the odds of having an ocular comorbidity if a subject reported moderate or worse difficulty on the activity adjusted for age, race, and sex..

Based on the preceding analyses, we concluded that the items recognising people, reading large print, and cooking contributed little to detection of functional impairment in cataract surgery candidates. Over 90% of cataract surgery patients had no difficulty preoperatively with these three tasks. In addition, less than 1% of the study population had moderate or greater difficulty on any of these three questions and did not report moderate or greater difficulty on two or more other questions. Finally, these three items showed poor correlation with the overall VF-14 score ($r < 0.4$) and had low effect sizes.

These items had originally been included in the VF-14 to provide an indication of disease severity. Those with severe visual impairment, it was believed, would be more likely to have difficulty on these relatively simple tasks.

Table 8 shows that the presence of one or more of three ocular comorbidities (diabetic retinopathy, macular degeneration, or glaucoma) was associated with an age adjusted threefold increased likelihood of having moderate or worse difficulty recognising people. This association with ocular comorbidity and difficulty recognising people was driven primarily by an association between difficulty recognising people and an association between difficulty with cooking and the presence of macular degeneration (OR = 3.3; 95% CI 1.6, 6.7) or glaucoma (OR = 2.0; 95% CI 0.9 to 4.2). Moderate or worse difficulty in cooking also was significantly associated with ocular comorbidity (OR = 1.9; 95% CI 1.1, 3.5), again largely due to macular degeneration (OR = 2.3; 95% CI 1.0, 5.0) or glaucoma (OR = 1.9; 95% CI 0.9, 4.1). The only other VF-14 item with a statistically significant correlation with the ocular comorbidity index was seeing steps or curbs (OR = 1.6; 95% CI 1.1, 2.4), which was due to increased difficulty with this item among individuals with glaucoma (OR = 1.9, 95% CI 1.1, 3.2).

**Table 9** Correlation between the VF-11, VF-14, and measures of trouble and satisfaction with vision (n = 551)*

| | Baseline trouble† | Baseline satisfaction† | Change in trouble‡ | Change in satisfaction‡ |
|---|---|---|---|---|
| VF-11 | –0.45 | 0.32 | –0.36 | 0.31 |
| VF-14 | –0.45 | 0.33 | –0.37 | 0.31 |

*Spearman correlation coefficients reported. The correlation between the preoperative VF-14 and the preoperative VF-11 is 0.99.
†Correlations are with the baseline VF-11 and VF-14.
‡Correlations are with the change in the VF-11 and the VF-14.

**Table 10** Summary of VF-14 items that performed poorly on various analyses

| Type of analysis | Specific analysis | Items that performed poorly |
|---|---|---|
| Low applicability | 1. ≥90% with no difficulty | Recognising people<br>Reading large print<br>Cooking |
| | 2. Rarely performed activity | Sports |
| Redundancy of item with other items | Pairwise correlations | None |
| Measuring something other than cataract | 1. Low correlations with remaining 13 items | Recognising people<br>Reading large print<br>Cooking<br>Night driving |
| | 2. Difficulty with activity not seen with other items | Night driving |
| | 3. Associated with ocular comorbidity | Recognising people<br>Cooking |
| Poor responsiveness to cataract surgery | 1. Correlation with change in item score and change in trouble and satisfaction | Recognising people<br>Reading large print |
| | 2. Effect size of the item (<0.3) | Recognising people<br>Reading large print<br>Cooking<br>Seeing steps or curbs |
| | 3. Effect size of VF-13 after removing | None |
| Poor screening performance of item | Other activities would identify moderate to severe difficulty | Recognising people<br>Reading large print<br>Cooking<br>Seeing steps or curbs<br>Writing checks<br>Playing games<br>Playing sports<br>Watching TV |

Removal of recognising people, reading large print, and cooking from the VF-14 resulted in a VF-11 with good internal consistency (Cronbach's $\alpha = 0.83$). The VF-11 also had a slightly larger effect size than the VF-14 (1.09 versus 0.99), as well as a comparable degree of correlation with change in trouble with vision and satisfaction with vision after cataract surgery (Table 9). The Spearman correlation between the VF-14 and VF-11 was 0.99 ($p < 0.0001$).

## DISCUSSION

Several studies have shown that functional status is more strongly correlated with self reported trouble and satisfaction with vision than is Snellen acuity,[1] at least in the range of vision loss represented among people presenting for cataract surgery.[5–7] This analysis was undertaken to assess the contribution of each item in one measure of functional impairment related to vision, the VF-14, as well as to evaluate the impact of deleting various items from the VF-14. Several of our findings provide a rationale for removing three items from the VF-14 (Table 10 summarises these). Fewer than 10% of patients awaiting cataract surgery who were enrolled in the Cataract PORT study had any difficulty recognising people when they are close, reading large print, or cooking. Less than 1% of the total population reported moderate or greater difficulty in one of these activities in the absence of comparable difficulty with another activity in the VF-14. These items are therefore plagued by a ceiling effect in which almost all subjects are at the top of the scale and therefore cannot improve with treatment. In addition, difficulty recognising people and difficulty cooking were both more strongly associated with ocular comorbidity than with cataract in our population. Removing recognising people, reading large print, and cooking from the VF-14 results in an 11 item index (the VF-11). This shorter index has good internal consistency, and is as strongly correlated as the VF-14 with preoperative trouble with vision and satisfaction with vision. In addition, the correlation between change in the VF-11 score after cataract surgery and changes in trouble with vision and satisfaction with vision were of a similar magnitude to those for the VF-14.

Responsiveness of individual items to cataract surgery was measured in a subset of subjects who only had one cataract removed at 4 months. These individuals were similar to those who had both cataracts removed in terms of age, race, and sex, but had higher VF-14 scores at baseline. It is possible that the responses of this subset to the VF-14 at 4 months may differ from that of all patients undergoing first eye cataract surgery. Those who elected not to have the second surgery performed may have been less satisfied with the surgery than those who chose to undergo a second operation. Conversely, those who were most satisfied may have elected to hold off on a second procedure. It is therefore difficult to predict how the bias introduced influenced the results.

Uusitalo and colleagues recently proposed removing seven items from the VF-14 to create a VF-7.[11] The authors relied solely on the correlation between change in individual items and change in patient satisfaction following cataract surgery to select items to remove from the VF-14. While they demonstrated good correlation between change in trouble and satisfaction after cataract surgery and change in the VF-7 score (as would be expected given the methodology used to develop the VF-7), they did not report on the internal consistency of the VF-7 in their patient population, or the likelihood that the VF-7 would fail to identify at least one type of functional impairment in patients undergoing cataract surgery. Our analysis found that the VF-11 could not be reduced further without jeopardising one or more important aspects of the performance of the index. Decreasing the number of questions posed to a patient reduces the time needed to use a standardised instrument and may increase the likelihood that clinicians will perform a systematic evaluation of functional

status. While we had hoped to be able to significantly shorten the VF-14, our analysis justified removing only three items. Removal of more items from the index would only weaken its utility in quantifying functional limitations caused by cataract. However, removing only three items will result in minimal time savings. Given the long track record of the VF-14, and the documented responsiveness and reliability of the instrument, we believe it is not advisable to remove items from this instrument which has been validated as a tool to measure disability related to cataract and corneal and retinal diseases.

## ACKNOWLEDGEMENTS

Funding: HS-06280; K-23 EY00358.

Dr Friedman also is funded by a the Robert E McCormick Scholar Award from Research to Prevent Blindness

The authors have no proprietary interest in the instrument evaluated in this paper.

### Authors' affiliations

**D S Friedman, J M Tielsch, S Vitale, O D Schein,** Dana Center for Preventive Ophthalmology, Wilmer Eye Institute, Johns Hopkins University, Baltimore, MD, USA
**E B Bass,** Departments of Medicine and Health Policy and Management, Johns Hopkins Medical Institutions, Baltimore, MD, USA
**E P Steinberg,** Resolution Health Strategies, Washington DC, USA

## REFERENCES

1 **Steinberg EP**, Tielsch JM, Schein OD, *et al*. The VF-14. *Arch Ophthalmol* 1994;**112**:630–8.
2 **Mangione CM**, Phillips RS, Seddon JM, *et al*. Development of the 'Activities of Daily Vision Scale'. *Med Care* 1992;**30**:1111–26.
3 **Lundstrom M**, Roos P, Jensen S, *et al*. Catquest questionnaire for use in cataract surgery care: description, validity, and reliability. *J Cataract Refract Surg* 1997;**23**:1226–36.
4 **Anderson G**, Black C, Dunn E, *et al*. Willingness to pay to shorten waiting time for cataract surgery. *Health Affairs* 1997;**16**:181–90.
5 **Steinberg EP**, Tielsch JM, Schein OD, *et al*. National study of cataract surgery outcomes. Variation in 4-month postoperative outcomes as reflected in multiple outcome measures. *Ophthalmology* 1994;**101**:1131–41.
6 **Alonso J**, Espallargues M, Andersen TF, *et al*. International applicability of the VF-14. *Ophthalmology* 1997;**104**:799–807.
7 **Mangione CM**, Orav EJ, Lawrence MG, *et al*. Prediction of visual function after cataract surgery. A prospectively validated model. *Arch Ophthalmol* 1995;**113**:1305–11.
8 **Schein OD**, Steinberg EP, Cassard SD, *et al*. Predictors of outcome in patients who underwent cataract surgery. *Ophthalmology* 1995;**102**:817–23.
9 **Agency for Health Care Policy and Research**. *Cataract in adults: management of functional impairment.* Rockville, MD, US DHHS, 1993.
10 **American Academy of Ophthalmology and American Society of Cataract and Refractive Surgery**. White paper on cataract surgery. *J Cataract Refractive Surg* 1996;**22**:645–50.
11 **Uusitalo RJ**, Brans T, Pessi T, *et al*. Evaluating cataract surgery gains by assessing patients' quality of life using the VF-7. *J Cataract Refract Surg* 1999;**24**:989–94.
12 **Cassard SD**, Patrick DL, Damiano AM, *et al*. Reproducibility and responsiveness of the VF-14. An index of functional impairment in patients with cataracts. *Arch Ophthalmol* 1995;**113**:1508–13.
13 **Deyo R**, Diehr P, Patrick D. Reproducibility and responsiveness of health status measures: statistics and strategies for evaluation. *Control Clin Trials* 1991;**12**:142S–58S.
14 **Kazis L**, Anderwon J, Meenan R. Effect sizes for interpreting changes in health status. *Med Care* 1989;**27(suppl)**:S178–89.
15 **Cohen J**. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
16 **Cronbach L**. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;**16**:297–334.